HW1   Name: Sergio Coronado   Date: 01/28/16

Goal: To Become acquainted with using python to parse data files and perform basic analysis operations as well as to create plots of data.

Approach: The python code was written by writing a class that would contain the parsed data. The data file to parse is passed through the program parameters followed by the name of a file to output the chart to. An example of the usage is as follows:

      $ ./Coronado.HW.1.py < data_file> < hist_output_file >

The libraries used for this program was matplotlib and statistics. These where used for plotting the data and performing basic statistical operations on the data. After, analyzing the individual packet distribution the data set was then looked at on a per IP Address Basis and finally in A per Subnet address basis.

Results: The data set in tcpset1.data contained 24999 entries with a minum packet size of zero and a maximum packet size of 1460. The packet size distribution had a mean size of 744.24 and a median size of 730. The Variance of the distribution was found to be 473,695. It was found that most of the packets tended to accumulate around the minimum and maximum as shown in figure 1.
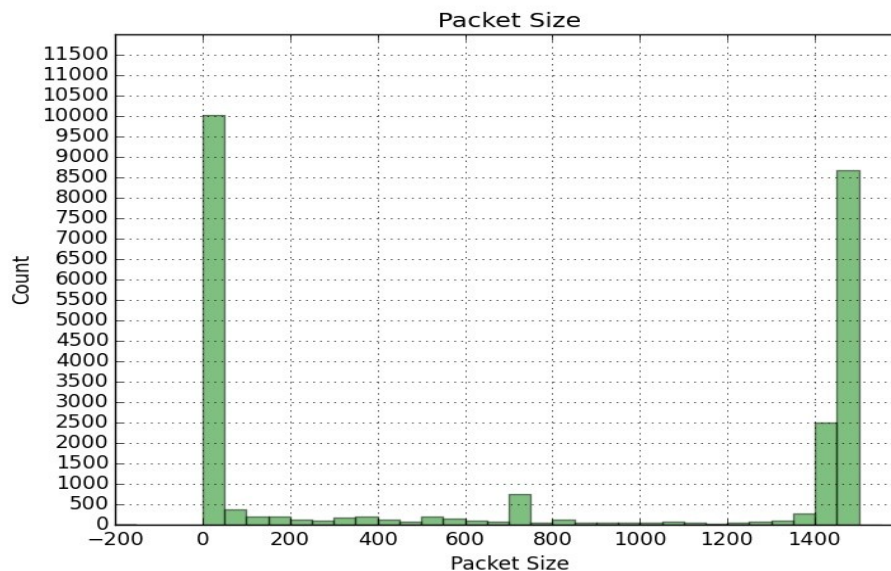

FIG 1: Histogram of packet sizes in tcpset1.data Interval of 200 bytes

When looking at the incoming and outgoing trafffic per IP the IP addres 8.3.209.233.80 seemed to send the most bytes and IP address 129.63.56.9.49730 received the most bytes. It was alsoe found that the subnet 129.63 was responsible for the most incoming and outgoing traffic.

References:
http://matplotlib.org/examples/
https://docs.python.org

Appendix A: Data Collected

**** Stats per Connection ****

Number of Entries: 24999
Max: 1460
Min: 0
Mean: 744.244649786
Median: 730
Mode: 0
Var: 473695.104877

**** Stats per Specific IP ****

Outgoing Min: 0
Outgoing Max: 954232
        IP: 8.3.209.233.80
Incoming Min: 0
Incoming Max: 954232
        IP: 129.63.56.9.49730
Mean Outgoing: 5792.45703611
Median Outgoing: 18.0
Variance Outgoing: 1298170660.18
Mean Incoming: 6066.30974894
Median Incoming: 56
Variance Incoming: 1198038342.87

**** Stats per Specific Subnet ****

Outgoing Min: 0
Outgoing Max: 4084215
        Subnet: 129.63
Incoming Min: 0
Incoming Max: 14541135
        Subnet: 129.63
Mean Outgoing: 17519.1826742
Median Outgoing: 117.0
Variance Outgoing: 21326567776.9
Mean Incoming: 16867.9709882
Median Incoming: 14
Variance Incoming: 1.91839151602e+11